



# Methods for Construction of *ab initio* Potential Energy Surfaces (PESs) Describing Large Amplitude Motion

---

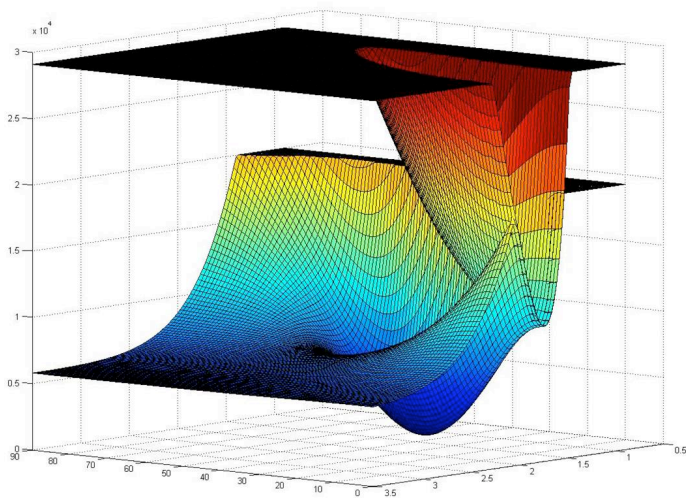
Richard Dawes and Ahren W. Jasper  
Combustion Research Facility-Sandia National Laboratories

2009 Symposium on Chemical Physics at the University of Waterloo





# Fitted Potential Energy Surfaces (PESs) via interpolation of *ab initio* data



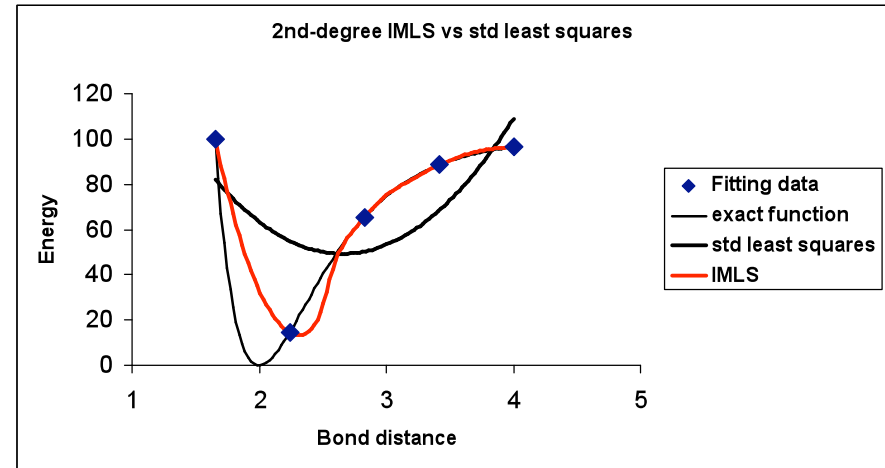
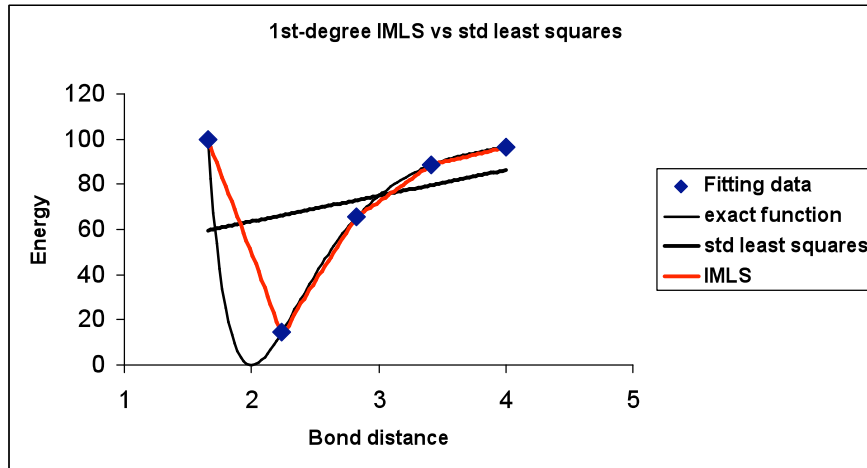
- Fitting methods are based on local Interpolative Moving Least Squares (L-IMLS)

JCP **130** 144107 (2009), JPC A **113**(16) 4626 (2009),  
JPC A **113**(16) 4709 (2009) , JCP **128** 084107 (2008)

- Interpolation between a number of stored local basis expansions
- Provides automated fitting. Runs in parallel on a computer cluster interfaced with popular electronic structure codes, adding data at automatically determined locations designed to rapidly converge the fit.
- Interpolates *through* data points making fit systematically improvable and ensuring correct degeneracy patterns on fitted PESs for multiple surfaces.



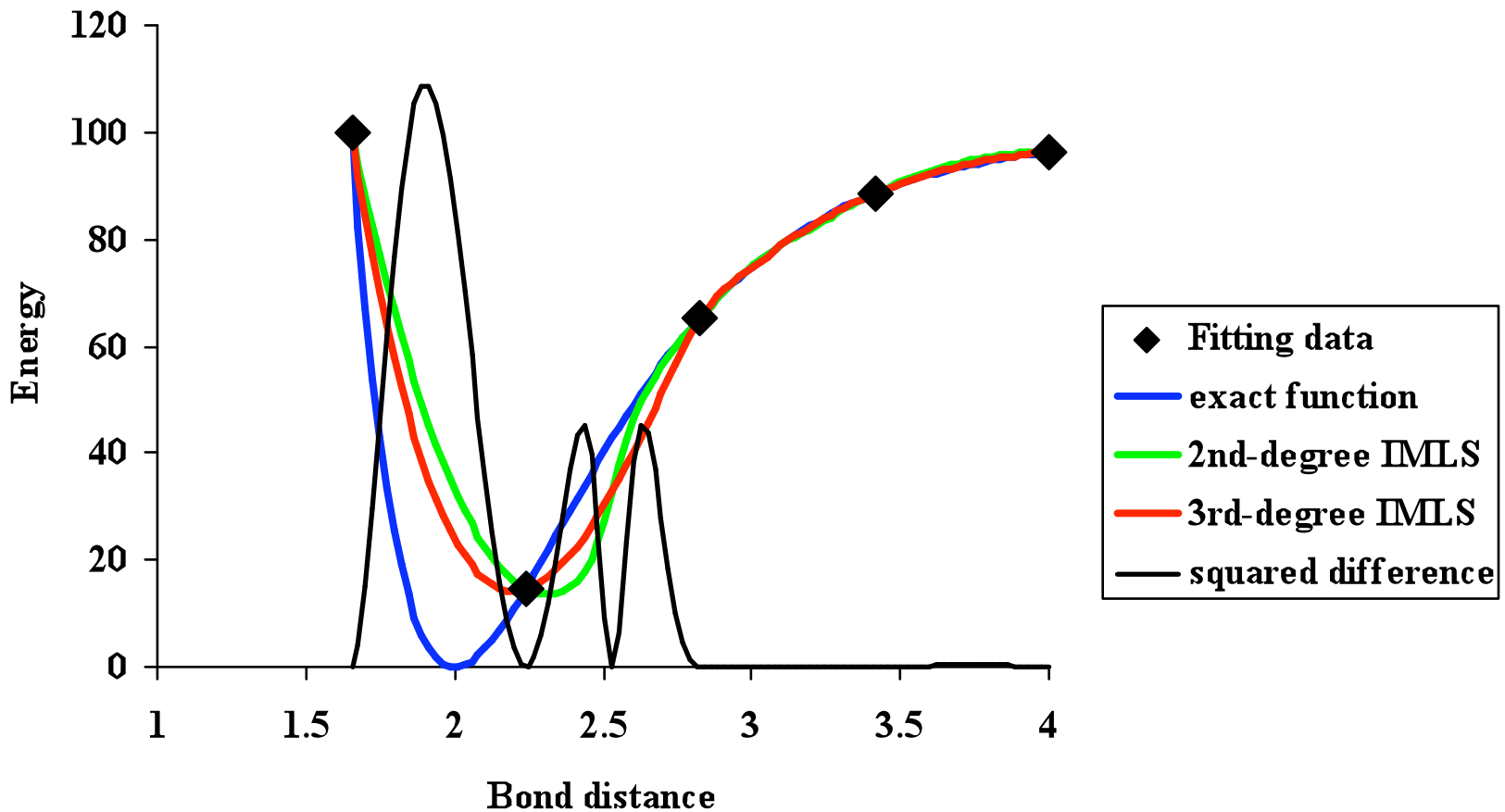
# IMLS



- Fit passes through data
- Much more flexible than a single expansion using the same basis
- Apply weights to robust linear algebra least squares solvers (e.g. LAPACK)

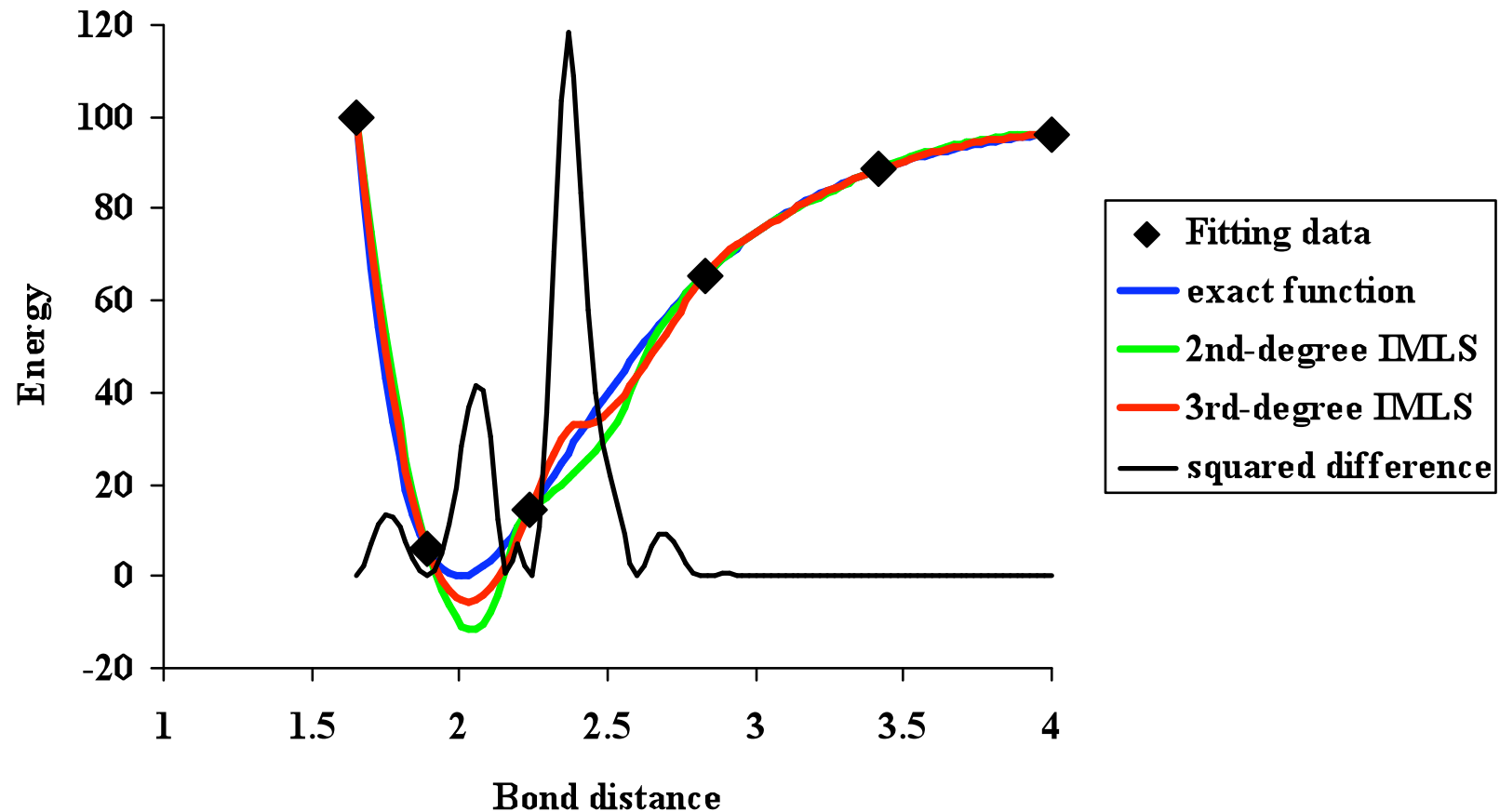


# Automatic PES generation: 5 seed points



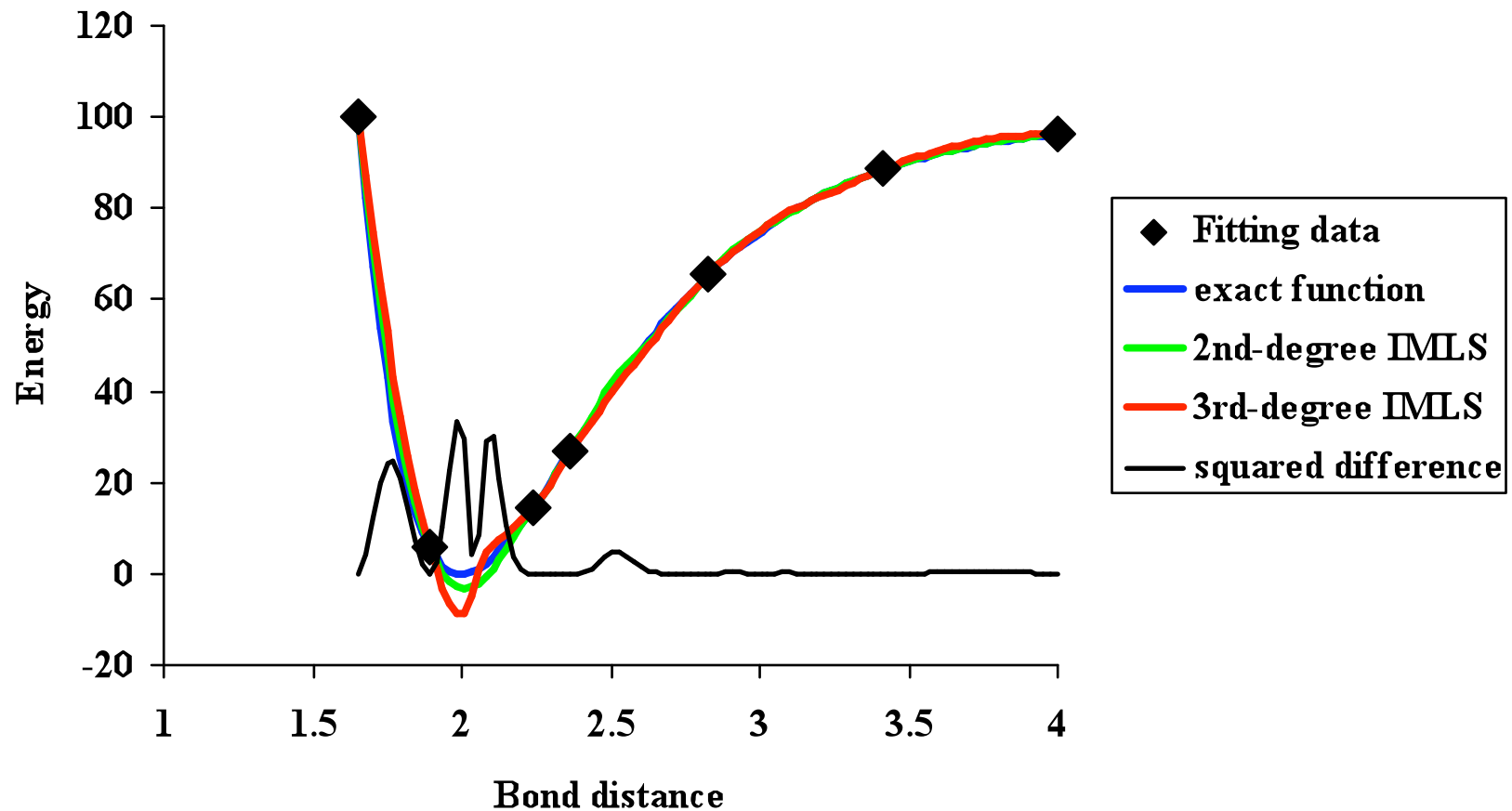


# Automatic PES generation: 5 seed points + 1 automatically generated point



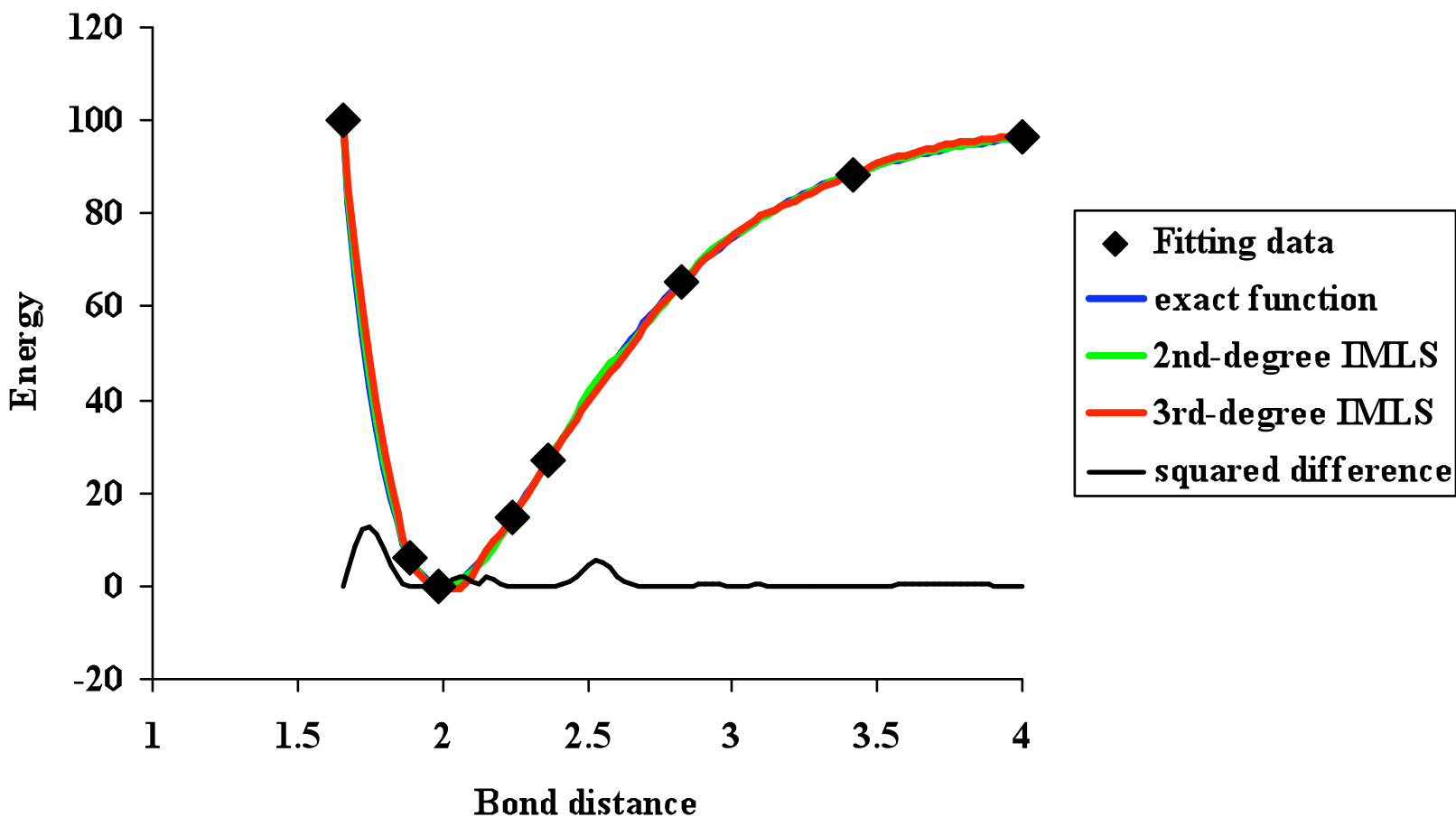


# Automatic PES generation: 5 seed points + 2 automatically generated points





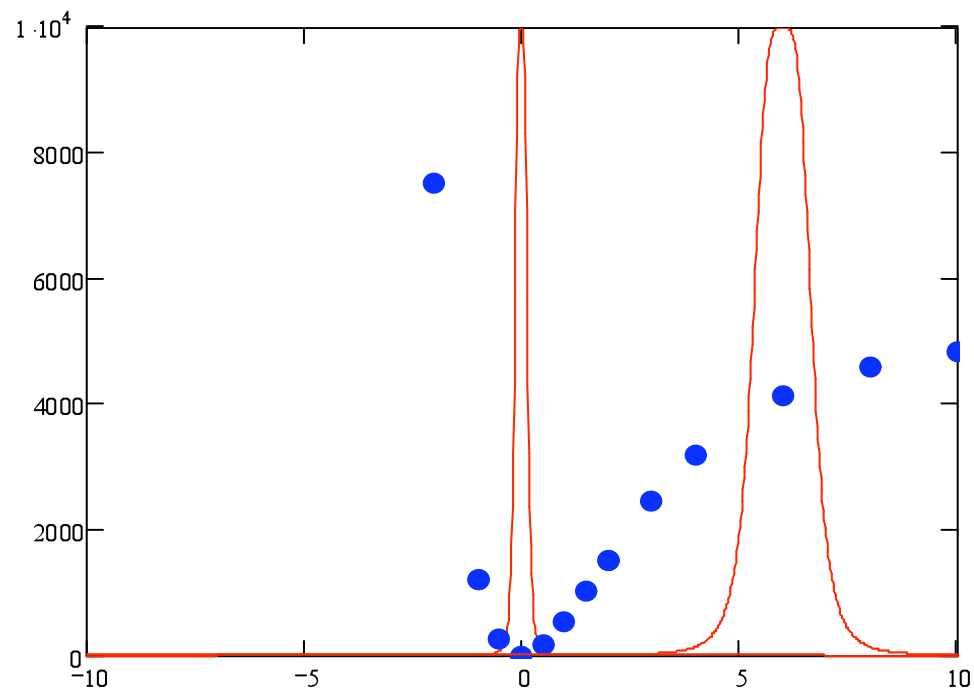
# Automatic PES generation: 5 seed points + 3 automatically generated points





# Density adaptive weight function

$$w_i(Z) = e^{\left(\frac{z-z_i}{d(i)}\right)^2} / \left(\left(\frac{z-z_i}{d(i)}\right)^{2p} + \varepsilon\right)$$





# Essential Components of IMLS-based Methods

- Choice of **coordinates** to describe system
- Fitting **basis**
- Well-defined **distance metric** between stored expansion points
- Interpolative **weight function**



# Applications to three-atom systems

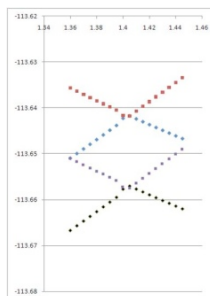
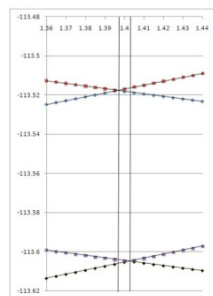
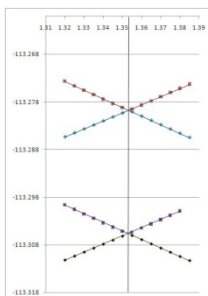
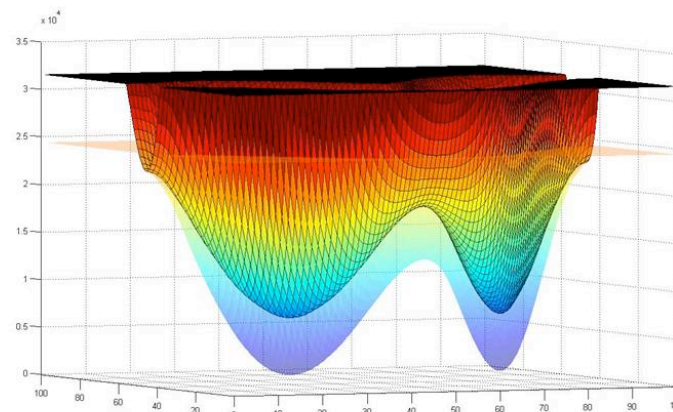
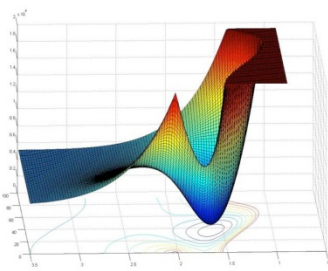
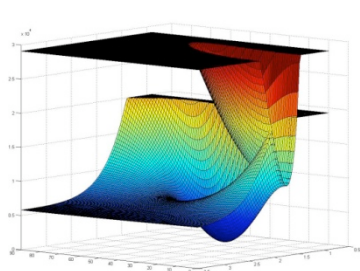
- Small 3-D configuration space
  - Coordinates and basis less important
    - Coordinates: valence, Jacobi, Radau or internuclear distances
    - Basis: expansion-centered polynomial in choice of coordinates
  - Simple distance metric
    - Generalized distance using differences in coordinates
- Extremely accurate and efficient fits obtained if *ab initio* method is “well-behaved”
  - Only a few hundred points are typically required to achieve wavenumber accurate fits for simple topologies
  - Allows direct evaluation of *ab initio* methods for prediction of spectroscopic levels
- CH<sub>2</sub> and HCN (JPC A **113**(16) 4709 (2009))
  - CCSD(T)/CBS, MRCI+Q/CBS+C-V+rel.+NBO...
  - 2-3 cm<sup>-1</sup> RMSE over large sets of expt. levels



# Dynamically-weighted state-averaged multireference electronic structure theory

$$E^{SA-MCSCF} = \sum_{i=0}^n w_i E_i \quad E^{DW-MCSCF} = \sum_{i=0}^n w_i (E_i - E_0) E_i$$

$$E^{GDW-MCSCF} = \sum_{i=0}^n w_i (|E_i - E_j|) E_i$$

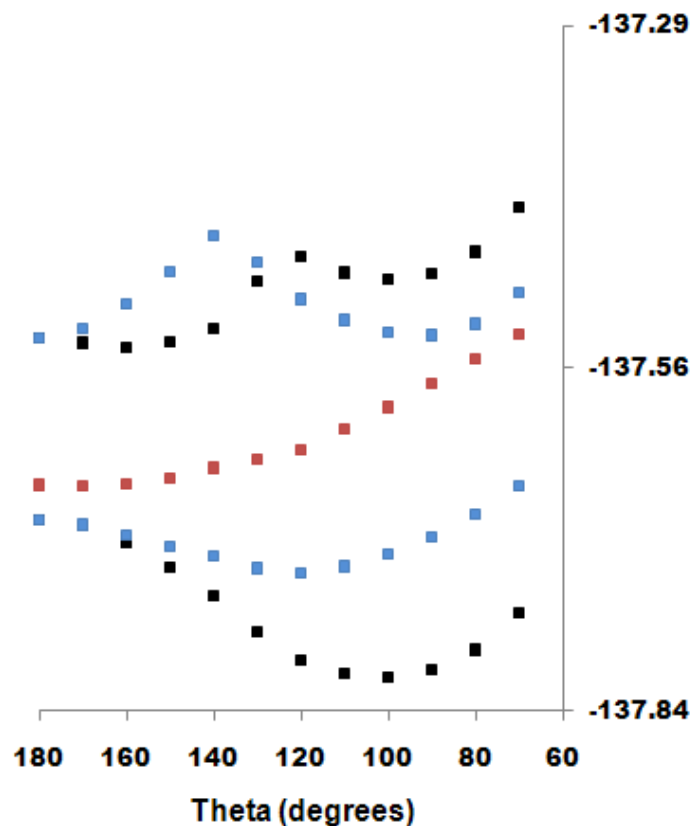


- Low-lying state(s)
- Good agreement with expt. ( $\sim 10 \text{ cm}^{-1}$ )
- Problems near CI
- Not a challenging test for GDW

M.P. Deskevich, D.J. Nesbitt and H-J. Werner, JCP **120** 7281 (2004)



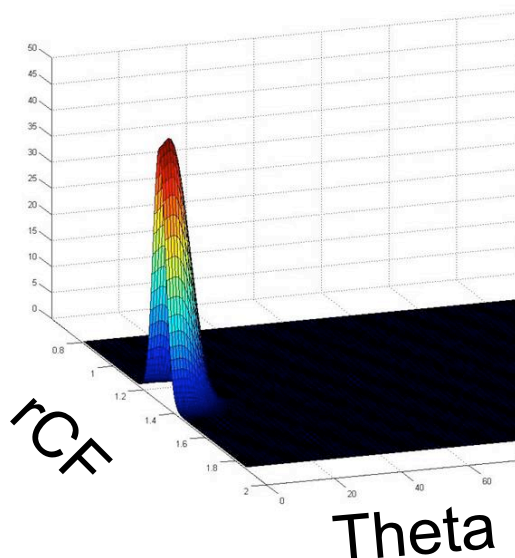
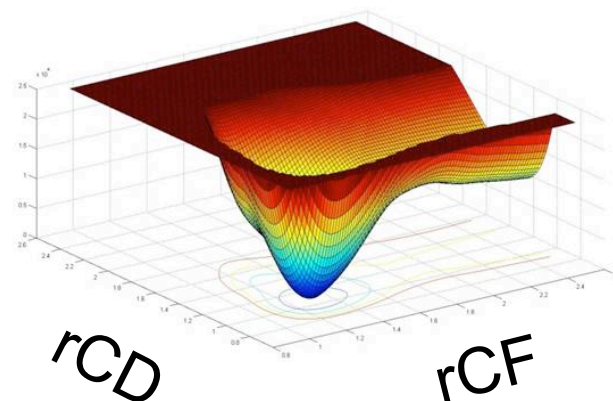
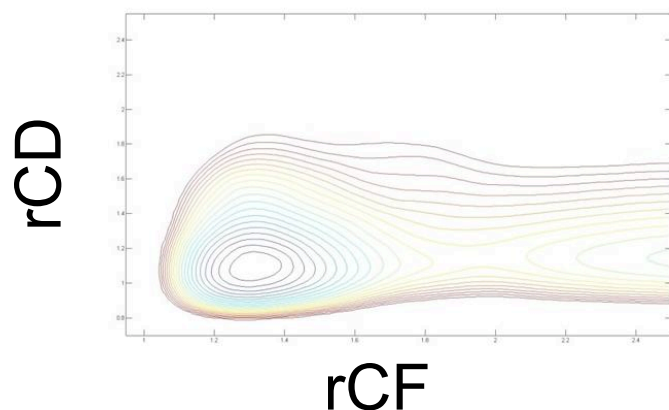
# $\tilde{B} (^1A')$ state of CDF



- Excellent test system for GDW scheme
- High-lying isolated state (31000 cm<sup>-1</sup>)
- Well characterized spectroscopically
- 2<sup>1</sup>A' state of interest (red)
- 1<sup>1</sup>A' and 3<sup>1</sup>A' states (black)
- 1<sup>1</sup>A'' and 2<sup>1</sup>A'' states (blue)
- Degeneracies at 180 degrees and related RT-coupling



# $\tilde{B} (^1A')$ state of CDF

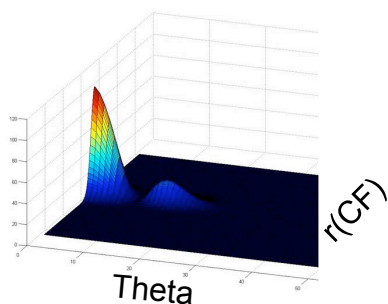
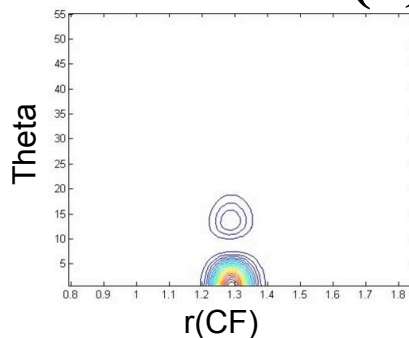


- Ground state: maximum probability in theta not quite at 0 degrees (linear)
- Barrier to linearity = 473 cm<sup>-1</sup>
- ZPE(CDF) = 2275 cm<sup>-1</sup>

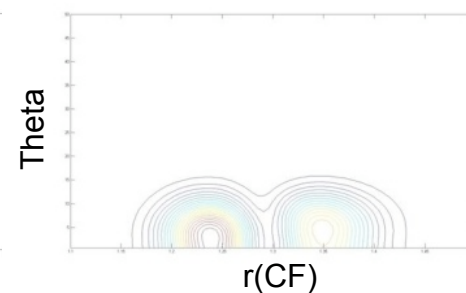
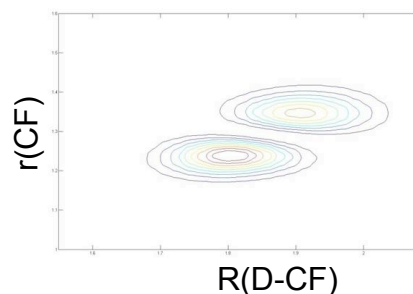


# $\tilde{B} (^1A')$ state of CDF

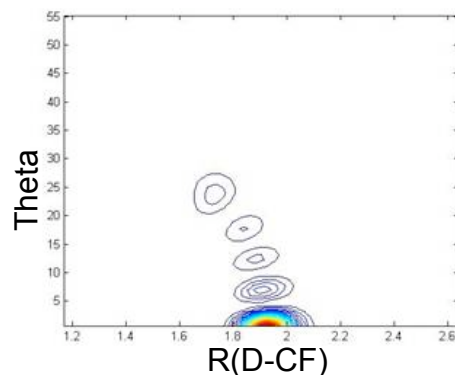
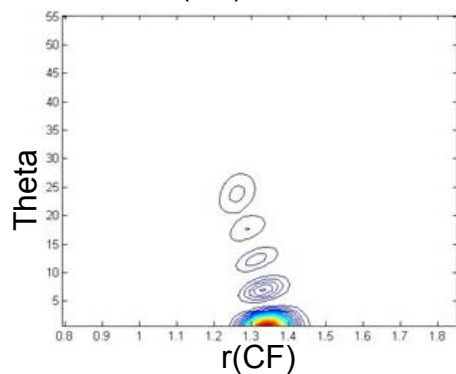
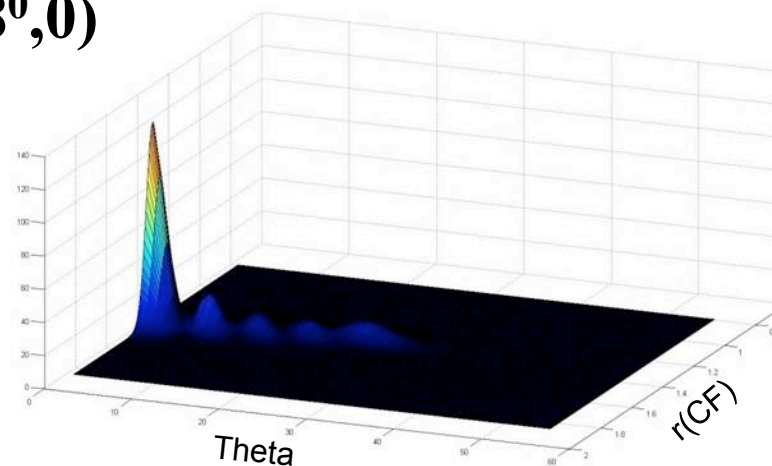
$(0,2^0,0)$



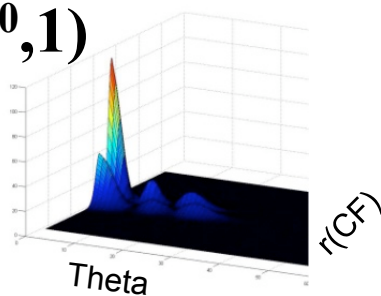
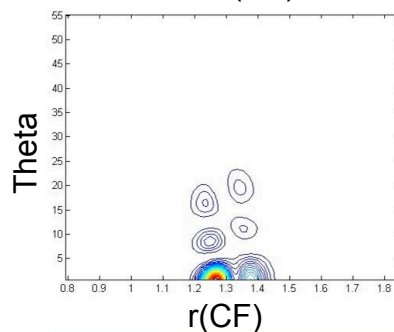
$(0,0^0,1)$



$(0,8^0,0)$



$(0,4^0,1)$





# $\tilde{B}$ ( $^1A'$ ) state of CDF

Level	EXPT (width)	GDW-MRCI/CBS	Error
(0,0 <sup>0</sup> ,0)	30782.4 (0.3)	---	
(0,2 <sup>0</sup> ,0)	904.5 (0.3)	900.9	-3.6
(0,4 <sup>0</sup> ,0)	1944.1 (1.1)	1930.5	-13.7
(0,6 <sup>0</sup> ,0)	3049.9 (2.6)	3037.1	-12.8
(0,8 <sup>0</sup> ,0)	4172.0 (3.7)	4155.2	-16.9
(0,10 <sup>0</sup> ,0)	5298.9 (4.3)	5313.2	14.3
(0,0 <sup>0</sup> ,1)	1292.4 (0.3)	1287.4	-5.0
(0,2 <sup>0</sup> ,1)	2182.7 (0.5)	2197.8	15.1
(0,4 <sup>0</sup> ,1)	3250.5 (1.6)	3244.1	-6.4
(0,6 <sup>0</sup> ,1)	4365.1 (3.6)	4358.7	-6.4
(0,8 <sup>0</sup> ,1)	5483.4 (5.0)	5483.7	0.4
(0,10 <sup>0</sup> ,1)	6605.7 (5.1)	6617.0	11.3
(0,2 <sup>0</sup> ,2)	3337.2 (0.8)	3335.8	-1.4
(0,4 <sup>0</sup> ,2)	4549.4 (4.3)	4553.2	3.8
(1,2 <sup>0</sup> ,0)	3316.5 (0.9)	3334.1	17.6
(1,4 <sup>0</sup> ,0)	4344.3 (3.1)	4347.5	3.2
MUSE			8.54



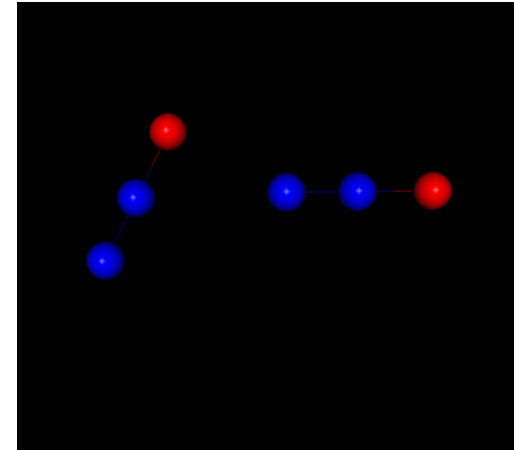
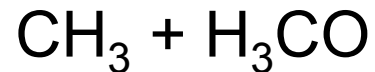
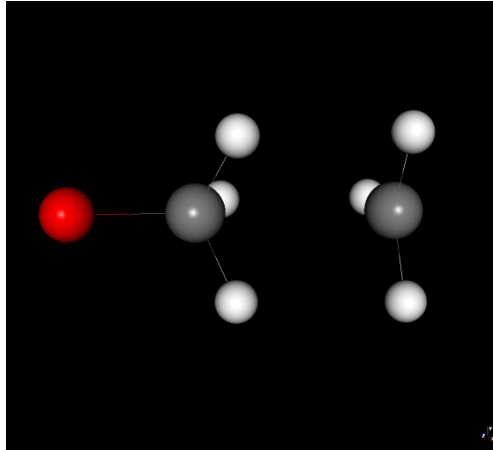
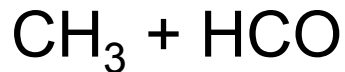
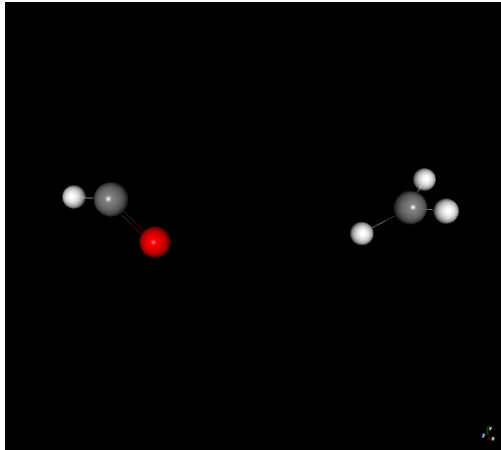
# Treating large amplitude motion

- Long range PESs with large amplitude motion
  - *E.g.* vdW or roaming molecular fragment systems
- 6-D large configuration space describes two general rigid fragments (4-D for two linear fragments)
- Orientational dependence is challenging for internuclear distance coordinates
- Relatively straightforward *ab initio*
- Develop specialized IMLS-based scheme considering three main components
  - Coordinates
  - Basis
  - Distance Metric





# Coordinates



- 6-D: center of mass distance, 5 Euler angles
  - $V(r, \gamma_1, \cos(\beta_1), \gamma_2, \cos(\beta_2), \alpha_1 - \alpha_2)$
- 4-D: center of mass distance, 2 angles and 1 torsion
  - $V(r, \cos(\theta_1), \cos(\theta_2), \varphi)$



# Basis

- Angles: Real rotation matrices based on Wigner rotation functions  $R_{m',m}^l(\alpha, \beta, \gamma)$

- Distance:  $\exp(\alpha r)$

- 6-D basis:

$$\sum_i \exp(\alpha r)^i \sum_{L_1, L_2=0}^{L \max} \sum_{M=0}^{\min(L_1, L_2)} \sum_{K_1=-L_1}^{L_1} \sum_{K_2=-L_2}^{L_2} R_{M, K_1}^{L_1}(0, \beta_1, \gamma_1) R_{-M, K_2}^{L_2}(-\alpha_{1-2}, \beta_2, \gamma_2)$$

- 4-D basis:

$$\sum_i \exp(\alpha r)^i \sum_{L_1, L_2=0}^{L \max} \sum_{(L_1+L_2 \leq L \max)}^{\min(L_1, L_2)} P_{L_1}^M(\cos(\theta_1)) P_{L_2}^M(\cos(\theta_2)) \cos(M\phi)$$



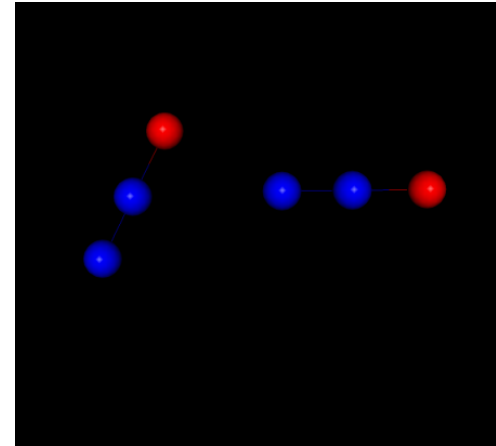
# Basis: dynamic conditioning

- Strategy to optimize fitting basis dynamically throughout configuration space
- SVD-based least squares solver
  - Near singular weight at data point to force interpolation
  - Large singular values correspond to well-determined linear combinations of the basis
  - Near-zero singular values correspond to poorly determined linear combinations of the basis
- Exclude linear combinations of basis (set condition number) until fit to neighbors is compromised  $>10\%$
- Improves fitting error to separate test set by  $> 30\%$
- Allows large general (even under-determined) basis



# Distance metric

- 4 required properties
  - $d(x,y) \geq 0$
  - $d(x,y) = 0$  iff  $x=y$
  - $d(x,y) = d(y,x)$
  - $d(x,z) \leq d(x,y) + d(y,z)$



- 4-D:  $d(x_1, x_2)^2 = c(r_1 - r_2)^2 + (\theta_{11} - \theta_{12})^2 + (\theta_{21} - \theta_{22})^2 + \sqrt{\sin(\theta_{11})\sin(\theta_{12})\sin(\theta_{21})\sin(\theta_{22})}(\phi_1 - \phi_2)^2$

- 6-D: Use Eckart-Sayvetz conditions to align Eckart frames of two structures (equivalent to minimizing sum of squared (mass-weighted) displacements)

$$\sum_{i=1}^N m_i \mathbf{r}_i' \times (\mathbf{U} \mathbf{r}_i) = 0$$

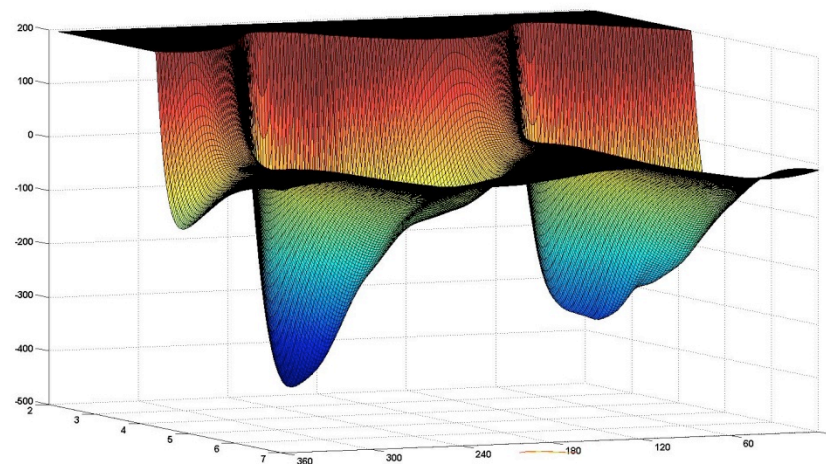
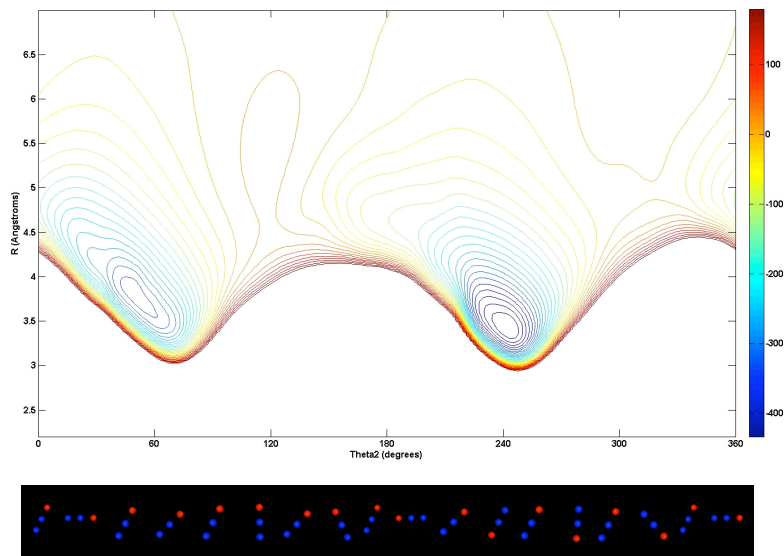
$$\min_{U \in SO(3)} \sum_{i=1}^N m_i |\mathbf{r}_i' - (\mathbf{U} \mathbf{r}_i)|^2$$

# Applications: (NNO)<sub>2</sub>

- *Ab initio*: CCSD(T), 44 electrons, no symmetry (in general)
  - G. Petersson, nZaP extrapolation bases
  - “Schwenke parameterize” CBS(2ZaP, 3ZaP)
  - Test estimated CBS at 5 points
  - Compare CP corrected
  - Simple 3ZaP more accurate than any 2-3 extrapolation scheme for 5 test points (3-4 CBS CP corrected most accurate 4-zeta scheme)
- Fit PES (10600 cm<sup>-1</sup> range) to 1-2 cm<sup>-1</sup> RMSE using <1600 single point energies
- An effective lmax of 30 was achieved using local expansions with lmax of only 6

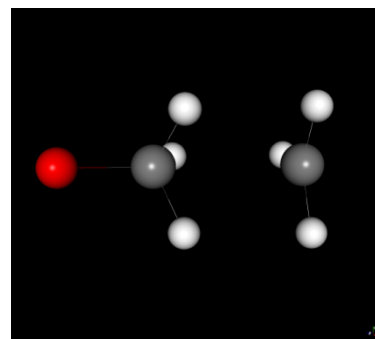
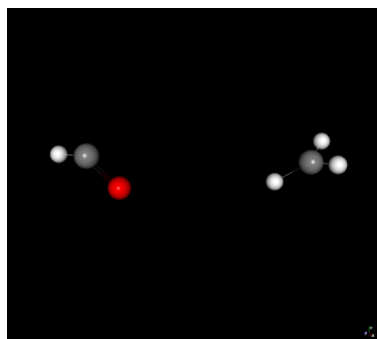


# Applications: $(\text{NNO})_2$



- Detailed PES characterization and analysis of ro-vibrational calculations underway

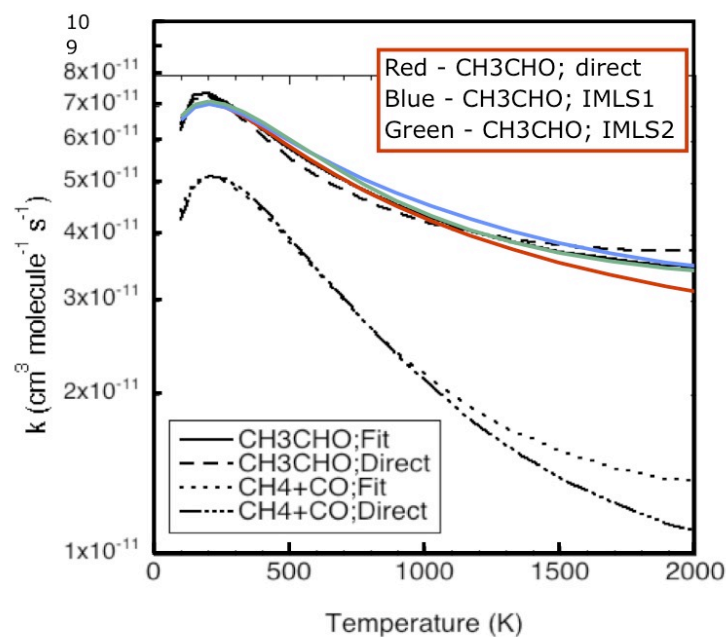
# Applications: Roaming radicals



- Acetaldehyde system previously fit by L. Harding using 8 internuclear-distance Morse-function expansions, switching functions and 103,000 *ab initio* points (CASPT2/avdz)
- Comparable quality of fit obtained with IMLS scheme using ~2000 *ab initio* points (energies and gradients)
- Application to series of related systems underway



# Applications: Roaming radicals



- Association kinetics on fitted surfaces (~2000 data points) agree with those computed using direct dynamics (~40000 *ab initio* calculations required)



# Summary

- IMLS scheme straightforward for high-accuracy in three atom systems
- Accurate and efficient representations of PESs for large configuration spaces require specialized:
  - Coordinates
  - Bases
  - Distance metrics
- Small number of *ab initio* data required enables rapid construction of PESs and/or the use of high-level electronic structure theory



# Acknowledgements

- Collaborators
  - Ro-vibrational calculations
    - Xiao-Gang Wang and Tucker Carrington Jr. (Queen's University)
  - Experimental spectroscopy (CHF)
    - Scott A. Reid (Marquette University)
  - Roaming radicals
    - Lawrence B. Harding and Stephen J. Klippenstein (Argonne National Labs)
  - IMLS fitting
    - Donald L. Thompson (Missouri) and Albert F. Wagner (Argonne)
- Funding
  - OBES DOE